

Page 1 of 7

## TECH CENTER 1600/2900

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/269,874A

DATE: 11/12/2001 TIME: 18:09:52

. . .

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\11122001\I269874A.raw

**ENTERED** 

```
4 <110> APPLICANT: Bujard, Hermann
 6 <120> TITLE OF INVENTION: Recombinant Process for Preparing a
 7
         Complete Malaria Antigen, GP190/MSP1
10 <130> FILE REFERENCE: GRUE-003
12 <140> CURRENT APPLICATION NUMBER: 09/269,874A
13 <141> CURRENT FILING DATE: 1999-08-02
15 <150> PRIOR APPLICATION NUMBER: PCT/EP97/05441
16 <151> PRIOR FILING DATE: 1997-10-02
18 <150> PRIOR APPLICATION NUMBER: DE 19640817.2
19 <151> PRIOR FILING DATE: 1996-10-02
21 <160> NUMBER OF SEQ ID NOS: 8
23 <170> SOFTWARE: FastSEQ for Windows Version 4.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 4920
                                    Λ
27 <212> TYPE: DNA
28 <213> ORGANISM: Plasmodium falciparum
30 <400> SEQUENCE: 1
31 atgaagatca tattetttt atgtteattt ettttttta ttataaatae acaatgtgta
                                                                         60
32 acacatgaaa gttatcaaga acttgtcaaa aaactagaag ctttagaaga tgcagtattg
                                                                        120
33 acaggttata gtttatttca aaaggaaaaa atggtattaa atgaaggaac aagtggaaca
                                                                        180
34 gctgttacaa ctagtacacc tggttcaaag ggttcagttg cttcaggtgg ttcaggtggc
                                                                        240
35 teagttgett caggtggete agttgettea ggtggeteag ttgetteagg tggeteagtt
                                                                        300
36 gcttcaggtg gttcaggtaa ttcaagacgt acaaatcctt cagataattc aagtgattca
                                                                        360
37 gatgctaaat cttacgctga tttaaaacac agagtacgaa attacttgtt aactatcaaa
                                                                        420
38 gaactcaaat atcctcaact ctttgattta actaatcata tgttaacttt gtgtgataat
                                                                        480
39 attcatggtt tcaaatattt aattgatgga tatgaagaaa ttaatgaatt attatataaa
                                                                        540
40 ttaaactttt attttgattt attaagagca aaattaaatg atgtatgtgc taatgattat
                                                                        600
41 tgtcaaatac ctttcaatct taaaattcgt gcaaatgaat tagacgtact taaaaaactt
                                                                        660
42 gtgttcggat atagaaaacc attagacaat attaaagata atgtaggaaa aatggaagat
                                                                        720
43 tacattaaaa aaaataaaaa aaccatagaa aatataaatg aattaattga agaaagtaag
                                                                        780
840
45 gctcaatatg atctttctat ttacaataaa caattagaag aagcacataa tttaataagc
                                                                        900
46 gttttagaaa aacgtattga cactttaaaa aaaaatgaaa acattaagga attacttgat
                                                                        960
47 aagataaatg aaattaaaaa tcccccaccg qccaattctg gaaatacacc aaatactctc
                                                                       1020
48 cttgataaga acaaaaaaat cgaggaacac gaaaaagaaa taaaagaaat tgccaaaact
                                                                       1080
49 attaaattta atattgatag tttatttact gatccacttg aattagaata ctatttaaga
                                                                       1140
50 qaaaaaaata aaaatattga tataagtgca aaggttgaaa caaaggaatc aactgaaccc
                                                                       1200
51 aatgaatate caaatggagt taettateet ttgteatata aegatattaa caatgettta
                                                                       1260
52 aatgaactta attettttgg tgatttaatt aatceatttg attatacaaa agaaceaagt
                                                                       1320
53 aaaaacatat atactgataa tgaaagaaaa aaattcataa atgaaattaa ggaaaaaatt
                                                                       1380
54 aaaatagaaa aaaaaaaaat tgaatctgat aaaaaatctt acgaagacag atctaagtct
                                                                       1440
55 ttaaatgata taacaaaaga atatgaaaaa ttacttaatg aaatttatga tagcaaattc
                                                                       1500
56 aataataata tagatttaac taatttcgaa aaaatgatgg gtaaaagata ttcatataaa
                                                                       156Ô
57 gttgagaaac ttacacacca taatactttt gcatcctatg aaaattctaa acataatctt
                                                                       1620
58 gaaaagttaa caaaagctct taaatatatg gaagattatt ctttaaggaa tatagtagtt
                                                                       1680
59 gaaaaagaat taaaatatta taaaaattta ataagcaaaa tagaaaatga gattgaaaca
                                                                       1740
60 ttagttgaaa atattaaaaa agatgaagaa cagctttttg aaaaaaaaat tactaaagac
                                                                       18Ŭ0
```

RAW SEQUENCE LISTING DATE: 11/12/2001 PATENT APPLICATION: US/09/269,874A TIME: 18:09:52

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\11122001\1269874A.raw

	-	cagatgaaaa	_	-	-	-	1860
	-	taatgaacaa		-	_		1920
		aacataatat	-		-		1980
		taattgtgtt					2040
		tgataaatga	-				2100
		caaccgaagg	_				2160
67	gcaggatctg	ctttagaagg	agattcagta	caagcacaag	cacaagaaca	aaaacaagca	2220
68	caaccaccag	taccagtacc	agtaccagaa	gcaaaagcac	aagtcccaac	accaccagca	2280
69	ccagtaaata	ataaaactga	aaatgtttcc	aaattagatt	atcttgaaaa	attatatgaa	2340
70	tttttaaata	cttcatatat	atgtcacaaa	tatattttgg	tttcacactc	aactatgaac	2400
71	gaaaagatat	taaaacaata	taaaattaca	aaggaggaag	aaagcaaatt	aagttcatgt	2460
72	gatccattag	acttattgtt	taatatacaa	aataacatac	ctgtaatgta	ttctatgttt	2520
	-	acaatagttt				-	2580
		ataaacttaa			_		2640
		catctgtaaa					2700
	-	aacccgaagt		_	•		2760
77		aattatttga				-	2820
	_	gtcaaaaaag	_				2880
		atgaatcttt		-	-		2940
		aatcaaaaag		_			3000
		catttgattt					3060
		cagttggtaa				_	3120
	-	caaaattgaa	_			_	3180
	-	acaaaaaaaa		-	-		3240
	-				-	•	
		tattgaaaca					3300
		ctttaagtga					3360
		tattaagtaa					3420
		catacttatc					3480
		aaaattatac				-	3540
		cttacaaaaa					3600
		ccgacacatt					3660
		ctaacacaat					3720
		ctatatttgg					3780
		cagtaactcc					3840
		tatatttaaa					3900
		ttatgacatt					3960
		atttcaaaaa					4020
		atgttgtcaa					4080
		gttataatta					4140
100	) gatgttcttg	, gatattataa	aatattatco	gaaaaatata	ı aatcagattt	agattcaatt	4200
						tttaaacaat	4260
102	2 attgagacct	tatataaaac	agttaatgat	: aaaattgatt	tatttgtaat	: tcatttagaa	4320
103	gcaaaagtto	taaattatac	atatgagaaa	ı tcaaacgtag	, aagttaaaat	aaaagaactt	4380
104	l aattacttaa	aaacaattca	agacaaattg	gcagatttta	aaaaaaataa	caatttcgtt	4440
						gttccttagt	4500
						: tgatggaaac	4560
						tccacaaaat	4620
						aaattacaaa	4680
						taatggtgga	4740
			-		-		-

RAW SEQUENCE LISTING DATE: 11/12/2001 PATENT APPLICATION: US/09/269,874A TIME: 18:09:52

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\11122001\1269874A.raw

		atgccaaatg					4800
		ctaaacctga					4860
	-	gaatatcatt	cttattaata	ctcatgttaa	tattatacag	tttcatttaa	4920
	<210> SEQ						
	<211> LENG'						
	<212> TYPE						
		NISM: Plasmo	odium falci	parum			
	<400> SEQU						
		tgaaaatcat		-	_		60
		cccacgaatc					120
		ccggatacag					180
		ccgttacaac					240
		ctgtggcctc					300
		caagcggcgg					360
		acgccaagtc		_			420
		agctgaagta					480
		ttcatggctt					540
		tgaatttcta					600
		gtcaaattcc					660
		tcttcggata					720
		atattaaaaa					780
		agaccataga					840
		cccagtacga					900
		tactggagaa		_		-	960
		agattaatga					1020
		tggacaagaa					1080
		ttaagttcaa	•	_	-		1140
		agaagaataa	-	_			1200
		atgaatatcc					1260
	-	acgagctcaa			_	-	1320
		agaatatcta					1380
		aaattgagaa					1440
		taaacgatat					1500
		acaataacat					1560
	-	tggagaaact	_	_	-	-	1620
		agaagctcac	-				1680
		agaaagaact					1740
		ttgttgagaa					1800
		aaaataaacc					1860
		aggtgctcct					1920
		tggagttaaa					1980
		catactacct					2040
		tcgagagcct					2100
		ccgagccttc					2160
		ccggttcagc					2220
		agcctccagt					2280
		ctgtgaataa					2340
		tcctgaatac					2400
161	actatgaacg	agaagattct	taaacagtac	aagataacca	aggaagagga	gagtaaactg	2460

RAW SEQUENCE LISTING DATE: 11/12/2001 PATENT APPLICATION: US/09/269,874A TIME: 18:09:52

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\11122001\I269874A.raw

```
162 tectettqtq atecaetqqa eetqetqtte aatatecaqa acaacattee eqttatqtat
                                                                          2520
163 totatgttcg atagcotcaa caattototo totoaactgt toatggagat atatgagaag
                                                                          2580
164 gagatggtet geaacetgta taaacteaaa gacaacgaca agattaagaa cettetggag
                                                                          2640
165 gaagctaaga aggteteeac etetgttaaa actetetett eeageteeat geaaceaetg
                                                                          2700
166 teteteacae etcaagacaa geeegaagtg agegetaaeg aegacaeete teaetegaee
                                                                          2760
167 aaccttaata actcactgaa actgtttgag aacatcctgt ctctcggcaa gaataagaac
                                                                          2820
168 atctaccaag aacttattgg acagaaatcg tccgagaact tctacgagaa gatactgaaa
                                                                          2880
169 gacagegaca cattetataa egagagette actaactteg tgaaatetaa ageegatgat
                                                                          2940
170 atcaactctc ttaacgatga atctaaacgt aagaagctgg aagaggacat caataagctg
                                                                          3000
171 aagaagacac tgcaactgag cttcgacctg tacaacaagt acaaactgaa actggagaga
                                                                          3060
172 ctcttcgaca agaagaagac agtcggcaag tataagatgc agatcaagaa gttgactctg
                                                                          3120
173 ctcaaggagc agcttgaaag caaactcaac tcactgaaca atccgaaaca cgtactgcag
                                                                          3180
174 aacttctcag tgttcttcaa caagaagaag gaagccgaga tcgccgagac agagaacact
                                                                          3240
175 ctggagaaca ccaagattct tctcaaacac tacaaaggcc tcgtcaagta ttataatggc
                                                                          3300
176 gagtettete etetgaagae teteteegag gagageatee agaeegagga taactaegee
                                                                          3360
177 agectegaga actteaaggt cetgtetaag etegaaggea agetgaagga caacetgaae
                                                                          3420
178 ctggagaaga agaagctcag ctacctctct agcggactgc atcacctgat cgccgagctc
                                                                          3480
179 aaggaagtca ttaagaacaa gaactacacc ggcaatagcc caagcgagaa taatacagac
                                                                          3540
180 gtgaataacg cactggaatc ttacaagaag ttcctgcctg aaggaacaga tgtcgccact
                                                                          3600
181 gtggtgtctg aatctggctc cgacacactg gagcagtctc aacctaagaa gcctgcatct
                                                                          3660
182 actcatgtcg gagccgagtc caatacaatt accacatctc agaacgtcga cgatgaggtc
                                                                          3720
183 gatgacgtca tcattgtgcc tatcttcggc gagagcgagg aggactacga tgacctcggc
                                                                          3780
184 caggtggtca ccggtgaggc tgtcactcct tccgtgattg ataacattct gtccaaaatc
                                                                          3840
185 gagaacgaat acgaagtgct ctatctgaaa cctctggcag gcgtctatag gtctctcaag
                                                                          3900
186 aaacagctgg agaataacgt gatgaccttc aatgtcaacg tgaaggacat tctgaacagc
                                                                          3960
187 cgctttaata agagagaaaa tttcaagaac gtcttggaga gcgacttgat tccctataaa
                                                                          4020
188 gacctgacct cctctaacta cgttgtcaag gacccataca agttcctcaa taaagagaag
                                                                          4080
189 agggataaat ttctgtctag ttacaactat atcaaggact ccatcgacac cgatatcaat
                                                                          4140
190 ttcgctaatg atgtgctggg gtattacaag atcctgagcg aaaaatacaa gtctgacctt
                                                                          4200
191 gactctatta aaaagtatat caacgataag caaggcgaga atgaaaaata tctgcccttc
                                                                          4260
192 ctgaataaca tcgaaaccct gtacaagaca gtgaacgaca aaatcgacct cttcgtaatt
                                                                          4320
193 cacctggagg ccaaggtcct caactatact tacgagaaga gcaatgtgga agttaaaatc
                                                                          4380
194 aaggagctga actacctcaa aacaatccaa gacaagctgg cagatttcaa gaaaaataac
                                                                          4440
195 aatttcgtcg gaattgcaga cctgtctacc gattataacc acaacaatct cctgaccaag
                                                                          4500
196 tttctgtcca ctggcatggt gttcgaaaac ctcgccaaaa cagtgctgag caatctgctc
                                                                          4560
197 gacggcaacc tgcagggcat gctgaacatc tcccagcacc aatgcgtgaa gaaacagtgc
                                                                          4620
198 ccccagaata gcggctgttt caggcatctg gacgagcgcg aagagtgcaa gtgtctcctg
                                                                          4680
199 aactacaaac aagaaggaga taagtgcgtg gagaacccaa accctacctg caatgaaaac
                                                                          4740
200 aatggcgggt gtgacgccga tgctaaatgc accgaggaag acagcggctc taacggaaag
                                                                          4800
201 aaaatcacat gegagtgtac taageeegae teetateeae tettegaegg gattttttge
                                                                          4860
202 tecageteta attteetggg cateteette etgetgatee teatgetgat eetgtacage
                                                                          4920
203 ttcatctaat agatcgatgg
                                                                          4940
205 <210> SEQ ID NO: 3
206 <211> LENGTH: 1639
207 <212> TYPE: PRT
208 <213> ORGANISM: Plasmodium falciparum
210 <400> SEQUENCE: 3
211 Met Lys Ile Ile Phe Phe Leu Cys Ser Phe Leu Phe Phe Ile Ile Asn
212 1
```

DATE: 11/12/2001 TIME: 18:09:52 RAW SEQUENCE LISTING PATENT APPLICATION: US/09/269,874A

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\11122001\I269874A.raw

213 214	Thr	Gln	Cys	Val 20	Thr	His	Glu	Ser	Tyr 25	Gln	Glu	Leu	Val	Lys 30	Lys	Leu
	Glu	Ala	Leu 35	Glu	Asp	Ala	Val	Leu 40		Gly	Tyr	Ser	Leu 45	-	Gln	Lys
217 218	Glu	Lys 50	Met	Val	Leu	Asn	Glu 55	Gly	Thr	Ser	Gly	Thr 60	Ala	Val	Thr	Thr
219 220		Thr	Pro	Gly	Ser	Lys 70	Gly	Ser	Val	Ala	Ser 75	Gly	Gly	Ser	Gly	Gly 80
221 222	Ser	Val	Ala	Ser	Gly 85	Gly	Ser	Val	Ala	Ser 90	Gly	Gly	Ser	Val	Ala 95	Ser
223 224	Gly	Gly	Ser	Val 100	Ala	Ser	Gly	Gly	Ser 105	Gly	Asn	Ser	Arg	Arg 110	Thr	Asn
225 226	Pro	Ser	Asp 115	Asn	Ser	Ser	Asp	Ser 120	Asp	Ala	Lys	Ser	Tyr 125	Ala	Asp	Leu
227 228	Lys	His 130	Arg	Val	Arg	Asn	Tyr 135	Leu	Leu	Thr	Ile	Lys 140	Glu	Leu	Lys	Tyr
	Pro 145	Gln	Leu	Phe	Asp	Leu 150	Thr	Asn	His	Met	Leu 155	Thr	Leu	Cys	Asp	Asn 160
232					165	Tyr				170					175	
234				180		Asn			185				_	190		
236			195			Asn		200					205			_
238		210				Leu	215			_	_	220			-	-
240	225					Asn 230					235	_	_			240
242					245	Lys				250					255	
244				260		Thr			265					270	_	
246			275			Leu		280					285			_
248		290				Glu -	295					300				_
250	305					Lys 310					315					320
252					325	Lys				330					335	
254				340		Asp			345					350		
256			355			Ala		360					365			
258		370				Glu	375					380				
260	385				•	Ala 390					395					400
2 U I	UDII	GIU	TAT	PTO	ASII	Gly	val	THE	тĂГ	LI.O	ren	ser	ryr	asn	ASP	тте

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/269,874A

DATE: 11/12/2001 TIME: 18:09:53

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\11122001\I269874A.raw